DOCUMENT RESUME

ED 054 914 RE 003 806

AUTHOR Esposito, Marietta Lee

TITLE Role of Immediate Recall in Selected Comprehension Tests at the Junior High Level.

PUB DATE Jun 71

NOTE 64p.; Doctoral dissertation submitted to Rutgers, The State University of New Jersey

EDRS PRICE MF-$0.65 HC-$3.29

DESCRIPTORS Analysis of Variance; Comprehension; *Correlation; Doctoral Theses; *Group Intelligence Tests; Intelligence Tests; *Junior High School Students; *Reading Comprehension; *Reading Research; Reading Tests; Recall (Psychological); Remedial Reading

ABSTRACT

An investigation was made to determine (1) if different reading comprehension subtests measure the same facets of comprehension, (2) what proportion of variance is accounted for in scores if verbal comprehension is partialed out, and (3) how highly group intelligence test scores correlate with comprehension tests which require immediate recall and those that do not. Forty seventh-grade remedial reading students were given the Spache Reading Scales and the Triggs Diagnostic Reading Test. Lorge-Thorndike group intelligence test scores were intercorrelated with the comprehension scores. Of the ten intercorrelations among the four comprehension subtests and IQ scores, four were significant at the .05 level or greater. The range of the partial correlations was from -.21 to .15. The variance accounted for by verbal comprehension ranged from -.04 to .03; other factors accounted for variation from .01 to .20. Triggs and IQ scores correlated highly at .50, but since a .39 correlation between IQ and Spache silent scores which demand immediate recall was also found, it was concluded that the third question was not satisfactorily answered. Suggestions for changes in the research design, a bibliography, appendixes, tables, and copies of the test are included. (BM)

ED054914

ROLE OF IMMEDIATE RECALL IN SELECTED COMPREHENSION TESTS AT THE JUNIOR HIGH SCHOOL LEVEL

A RESEARCH PAPER

SUBMITTED TO THE FACULTY

OF THE GRADUATE SCHOOL OF EDUCATION

OF

RUTGERS UNIVERSITY

THE STATE UNIVERSITY OF NEW JERSEY

BY

MARIETTA LEE ESPOSITO

IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE

OF

DOCTOR OF EDUCATION

NEW BRUNSWICK, NEW JERSEY JUNE, 1971

APPROVED: Marietta Klug

Eva Mountain

Ernest Haskington

DEAN: _____

RE003 806

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ABSTRACT

Purpose

The purpose of this study was to investigate: (1) if different comprehension subtests, as found in certain oral and silent reading tests, measure the same facets of comprehension; (2) what proportion of the variance is accounted for in comprehension test scores if the need for verbal comprehension of the material is partialed out of the score; and (3) how highly do group intelligence tests correlate with comprehension tests which necessitate immediate recall and those which do not.

Procedure

Five scores were collected for 40 seventh-grade remedial reading students. The Spache Reading Scales were administered first to the subjects in one sitting. The test was given to each student individually, and scores were collected for the oral and silent reading comprehension sections. The Triggs Diagnostic Reading Test was administered to subjects in groups of three or four. Scores were collected for the two comprehension subtests. The Lorge-Thorndike group intelligence test scores were obtained from school records.

Intercorrelations were computed for the four-reading comprehension subtests and intelligence test scores. Partial correlations were calculated, partialing out verbal comprehension.

Results

Three correlations were significant at the .01 level: Triggs Comprehension and IQ scores, Spache Oral and Spache Silent reading subtests, and Spache Silent and IQ scores. Correlations significant at the .05 level were: Triggs Story Comprehension and Triggs Comprehension scores, and Triggs Story Comprehension and Spache Silent Comprehension scores. A range of .01 percent to 25 percent of the total variance was accounted for in the intercorrelations.

The range of the partial correlations was from -.21 to .15. The amount of variance accounted for by variance comprehension ranged from -.04 to .03. The amount of variance accounted for by factors other than verbal comprehension was from .01 to .20.

Conclusions

Readers may use differing comprehension factors in seemingly similar circumstances. One of these variables--such as verbal comprehension--can be measured independently, and does account for some of the predictability of the response.

Comprehension is not, on the whole, general; and although there may be some overlap in abilities, different facets or skills are applied in various seemingly similar tests.

TABLE OF CONTENTS

	Page
LIST OF TABLES	iv
Chapter	
I. INTRODUCTION	1
The Problem	2
Statement of the Problem	2
Importance of the Study	3
Limitations of the Study	4
Definitions of Terms Used	5
II. REVIEW OF THE LITERATURE	6
Modal Differences of Silent and Oral Reading	6
Interpretation of Silent and Oral Reading Test Results	10
Nature(s) of Comprehension	15
Correlations of Intelligence and Reading Tests	21
Immediate Recall and Comprehension	24
III. PROCEDURES	27
Description of the Sample	27
Materials Used for the Study	28
Procedure of Testing	29
Treatment of the Data	31

TABLE OF CONTENTS (continued)

Chapter		Page
IV.	RESULTS AND DISCUSSION	34
V.	SUMMARY AND CONCLUSIONS	49
BIBLIOGRAPHY		53
APPENDIXES		
A.	Spache Diagnostic Reading Scales	57
B.	Triggs Diagnostic Reading Test	59
C.	Raw Data Scores	60

Appendixes A and B contain copyrighted material,
which is omitted.

LIST OF TABLES

Table		Page
1.	Means and Standard Deviations of Spache Diagnostic Scale, Triggs Diagnostic Test, and Lorge-Thorndike IQ (in Raw Score Form) . .	35
2.	Correlations among Spache Diagnostic Scales, Triggs Diagnostic Test, and Lorge-Thorndike IQ (in Raw Score Form)	37
3.	Proportional Overlap among Variables in Terms of Variance	39
4.	Partial Correlation Coefficients Eliminating the Triggs Comprehension Test (Verbal Comprehension)	41
5.	Amount of Variance Due to the Effect of Verbal Comprehension	41
6.	Amount of Variance After the Elimination of the Effect of Verbal Comprehension	42

CHAPTER I

INTRODUCTION

Recent years have brought about an increase in the number of remedial reading classes, the concentration of the growth being found at the intermediate and high school levels. These supplemental classes have as their major, generalized function the diagnosis and remediation of students' problems which are directly or indirectly related to reading.

The vertical and horizontal development of the role of diagnosis has brought about the need for more accurate and comprehensive diagnostic tools. The process of evaluation should enable the teacher: to ascertain the child's present level of ability and attainment; to estimate the amount, rate, and quality of learning; and to offer effective guidance and direction for future growth.

At one time educators accepted, with little question, the results of standardized tests as valid indicators of reading status. Accordingly, initial and final test scores were compared and gains in reading were estimated. As investigators began to question the validity of standardized testing, and as the concept of the reading

process became more and more comprehensive, this simple practice fell into disfavor. Today, most authorities recognize standardized test scores as partial and incomplete indicators of reading status.

Demand for practicality and thoroughness of the diagnostic tools has led to studies of the intrinsicality(s) of silent and oral reading and comparisons of the validities and nature of results of such tests.

Furthermore, the problem of how to measure comprehension in reading seems to be of major importance for two reasons: (1) to determine whether a student understands what is read, and (2) to know what to teach so that comprehension may be improved. In this era of great emphasis on reading problems, attention needs to be turned toward improved measurement of reading comprehension.

The Problem

Statement of the Problem

This study will investigate the following questions: (1) Do different comprehension subtests, as found in certain oral and silent reading tests, measure the same facets of comprehension? (2) What proportions of the variance is accounted for in comprehension test scores if the need for verbal comprehension of the material is partialed out of the score? (3) How highly do group intelligence tests correlate with comprehension tests

which necessitate immediate recall and those which do not?

Importance of the Study

This study developed from the observation that junior high school remedial reading teachers indiscriminately use oral and silent reading tests to measure students' present and potential levels of comprehension. This use of both types of tests is based on the assumption that both types measure comprehension equally well. Utilizing both types of evaluation devices in a remedial reading class, the investigator noted discrepancies between: (1) the students' abilities and the levels indicated by the tests; and/or (2) the areas of strengths and weaknesses in comprehension skills as measured by the tests and the students' actual performances. It was also noted that some of the tests required the reader to use much memory in order to carry out the comprehension questions.

There is a lack of research on the validity of comprehension measures used for the evaluation of remedial readers at the junior high school level. The validity attributed to these subtests at this level of instruction has developed through an upward extension of that validity as utilized in elementary school testing. This research study will provide specific data on whether these

comprehension subtests are appropriate in junior high school remedial reading evaluation procedures.

A review of the literature failed to uncover any study or research which sought to determine the amount of influence which immediate recall has upon measured comprehension. This study will extend the area of comprehension in the mentioned direction.

Limitations of the Study

The subtests of the tests utilized in the experiment were limited to those subtests labeled "comprehension."

At the junior high school level, the vocabulary levels of the Spache Diagnostic Scales are limited to children whose deficient reading ability will not allow them to exceed the 6.5 grade level ceiling set by the scales.

All but one of the students participating in this study succeeded in reaching this level, allowing the assumption that these 40 pupils might have also succeeded on more difficult vocabulary lists.

However, the lists would be of use if the evaluation involved another group of students, whose vocabulary recognition skills were less developed.

Definitions of Terms Used

Comprehension. Comprehension as used in this study was the number of correct answers the children obtained on questions about the story.

Instructional level. This term is used to designate the level of reading which most teachers would find acceptable in group or classroom practice.

Independent level. That grade level of supplementary instructional and recreational reading materials which the pupil can read to himself with adequate comprehension, even though he may experience some word-recognition difficulty.

Immediate recall. Immediate recall, as used in this study, will be the ability to answer questions pertaining to a read selection, using memory only, without benefit of referral back to the selection.

CHAPTER II

REVIEW OF THE LITERATURE

Modal Differences of Oral and Silent Reading

The literature indicated that: (1) administration of an oral and a silent reading test upon the same subject may produce results indicating varying reading levels, and/or (2) each type of test may be evaluating different domains of the reading process.

Fairbanks (1937) stated that the testing of oral reading is a good measure of an individual's reading processes, thus signifying his conclusion that poor silent readers tend to be poor oral readers in elements common to both types.

Swanson (1937) conducted experiments with college freshman calculating oral reading errors made when reading a paragraph. His premise was that if errors made orally are caused by perception difficulties, the same errors would occur during silent reading. A similar experiment was conducted with adults whose oral reading errors were calculated from a silent reading test as well as with comprehension of oral reading. There was a high correlation between the total number of errors under both conditions;

but correlations between oral reading errors and tests of silent reading comprehension were not high enough to permit predictions of individual performance. Swanson reasoned that this low relationship was due to psychological functions which are also influential on the acquisition of meaning and comprehension as well as perceptual accuracy.

An indication that frequency of oral inaccuracies tends to increase with decreased functioning of more complex psychological processes was found when the poor readers were not required to answer comprehension questions after their reading. On the other hand, the requirement to meet specific comprehension demands tended to produce more meaningful and consequently more accurate reading.

Swanson explained the results of his experiments as supportive of his premise that oral and silent reading are one and the same process.

Usually, the most vigorous defender of the two skills being the same in nature will make some concession as to the presence of some peripheral differences between oral and silent reading.

The basic difference, according to Dolch (1955), is not if a sound is made or not made (silent reading), but in the speed of the reading operation. Making noise, he proposed, does not create reading. Since full comprehension can be obtained at various rates, reading should

be studied or judged with various rates.

Other investigators have been convinced that there are more than slight differences between the two modes.

Gates (1947) considered effective oral reading more difficult than silent reading because it not only presupposes the ability to comprehend the material, but also involves other difficult-to-acquire abilities and aptitudes.

In the process of good oral reading, the eyes lead by a considerable distance the words being spoken. The better the oral reader, the larger this lead or eye-voice span is likely to be. A reasonable interpretation of this phenomenon seems to be that the pupil is attempting to anticipate the sequential drift of the author's thought in order to guide his own oral expression (Buswell, 1920; Fairbanks, 1937).

Anderson and Swanson (1937), Fairbanks (1937), and Swanson (1937) considered the dominant influence in effective reading, whether oral or silent, to be the central processes, including purpose and understanding--as opposed to the peripheral processes, such as eye movement. Eye movements in silent reading have been shown to be modified by: (1) changes in purpose, attitudes, and mental set, and (2) changes in the nature and difficulty of the materials.

The eye movements of good readers were remarkably

flexible in meeting changes in conditions. Moreover, in comparing the eye movements of identical pupils, the number and length of fixations and the number of the regressions were greater in oral than in silent reading. This seemed to indicate some adaptation of eye movements to the demands of the oral reading situation.

Earlier studies of O'Brien (1921) and Cole (1938) indicated that the average number of fixations in oral reading was greater than in silent reading; regressive movements were slightly more frequent; and the length of fixations was, on the average, greater in oral reading than in silent.

A commonly used measure of speed of perception is the duration of fixation pauses in reading. According to Schmidt (1917), the average duration of pauses is slightly over three-tenths of a second in silent reading, and slightly under four-tenths of a second in oral reading. The range of the number of regressions per line varied in silent reading from 4.0 to none, and in oral reading from 3.7 to 0.2.

The issue between oral and silent reading disappears, suggests Hildreth (1949), when initial reading lessons are based on the child's experiences, expressed in his own language. In these first lessons in associating meaning with print, oral expression is fundamental.

Silent "looking" and thinking are required in the same lesson so that the child is forced to concentrate on meaning and not merely on articulating a succession of speech sounds.

Generalizing the conclusions reached by researchers, oral reading and silent reading are essentially the same processes, utilizing similar aptitudes. The variation between the two modes may be due to physiological limitations set by the mode (e.g., length of fixation pause), the nature of materials, and motivations of the reader.

Of the two modes, the concensus was that oral reading is the more difficult of the two modes because each word must be distinctly pronounced as the reader holds the meaning of the passage in mind. The task is easier only to the extent that the reader is familiar with the vocabulary and context.

Interpretation of Silent and Oral Reading Test Results

Betts (1947) and Chall (1958)--sampling fifth graders--and McCracken (1964)--sampling sixth graders--compared equivalent scores of several standardized reading tests with the performance of pupils in functional reading materials. They concluded that: (1) standardized tests were not adequate for determining the level of achievement

of pupils at the lower or upper ends of the distribution; and (2) the standardized reading test scores were significantly higher than their reading performance as determined by an informal reading inventory.

The results of the studies of Julitta (1957) and Killgallon (1957) were supportive of the above. They concluded that the scores on standardized reading achievement tests placed children an average of one grade above their instructional level.

In his investigation of the relationship between the standardized test and informal estimates of reading levels among 1,400 pupils in grades 2 through 6, Botel (1957, p. 441) found:

1. In grade 2, 85 percent of the pupils were overrated by the standardized test from one to five levels. Eleven percent of the pupils were rated properly; 4 percent were underrated from one to two levels.

2. In grade 3, 68 percent of the pupils were overrated by the standardized test from one to five grades. Seventeen percent of the pupils were rated properly; 15 percent were underrated from one to three levels.

3. In the intermediate grades, on the average, about one-third of the pupils were overrated from one to five levels; one-third were rated properly; and one-third were underrated by the standardized test.

Botel claimed that the above evidence supported his view that there is little justification for using grade scores from standardized reading tests to place students in basal readers. He pointed out the complex relationship between a score on a test and the instructional level of a pupil; and that it is not a simple matter of adding or subtracting a constant figure from a pupil's reading score to arrive at the instructional reading level.

A study by Betts (1956), utilizing several standardized reading tests at the fifth-grade level, revealed that none of the tests was adequate for determining the level of students at extreme ends of the scales. Although 10 percent of the class did not exhibit desirable reading behavior on first-grade materials, some of the tests graded these pupils no lower than second-, third-, or fourth-grade level. In general, Betts concluded, standardized tests may be expected to rate those pupils from one to four grades above their manifest achievement levels. His findings substantiated Chall's (1958) analysis of standardized reading tests.

Harris (1953a) recommended that instruction begin one or two years below the grade level indicated by the test score on standardized silent reading tests.

Chall (1958) challenged the practice of placing

students in reading on a level below the grade scores as indicated on a standardized reading test. She believed that the scores obtained from these tests produced by children who lack confidence or who read very slowly may represent minimal estimates of performance. This type of child would benefit from a higher level of material.

Harris' (1956) view approached the middle ground: he asserted that formal reading test scores do not discriminate among the various reading levels. He stated that, although in most instances standardized scores generally reflect the instructional level, it must be pointed out that the reading performance of pupils who find the test materials too difficult, or who mostly guess on a standardized test, may yield a score indicating their frustrational level in reading.

Extending his median approach, Harris stated that oral examination has several advantages: (1) the question can allow freedom of response; (2) incorrect interpretation on the student's part can be corrected by the teacher; and (3) social communication between pupil and examiner is allowed. The major disadvantage, according to Harris, is that the subject is allotted only one answer.

It was found in clinics that, after remedial instruction, many fifth- and sixth-year children were able to reach their grade levels on untimed, or generously

timed, tests, whereas they frequently were unable to achieve more than third- or fourth-grade scores in short timed tests. Speed of reaction was being tested as well as speed of reading. The remedial reader needed time to organize himself (Newman, 1969).

Noting the level of difficulty of the hardest item done with clear success on a standardized test is one way, McCullough (1953) suggested, of determining the pupil's comfortable reading level.

Harris (1953b) suggested that one can make good use of standardized reading tests (silent) by analyzing the comprehension problems. But he cautions that discriminations must be used as one interprets the age, grade, or percentile scores. Individual understanding of the test directions may vary the way the student may work with them, thus lessening the effects of standardization. The student's need for surity of answers, or quantity of answers (regardless of correctness), may influence and change the scoring.

A majority of the researchers have concurred that standardized reading tests yield significantly higher or lower scores than do informal reading inventories (as characterized in this study by the Spache Diagnostic Reading Scales), thus indicating little justification for using these scores to classify children's reading levels.

Moreover, most researchers agreed that standardized tests were not adequate for determining students' reading levels at the extreme ends of the scales, often presenting distorted pictures of the achievement of the readers.

Nature(s) of Comprehension

The fact that the reading act cannot take place unless meaning is attached to what has been read has long been recognized in the field of reading education. Yet studies of reading comprehension have lagged behind other areas of testing reading. One of the reasons for this lack lies in the nature of the complexity of the activity, for its performance is usually less overt and much has to be assessed indirectly by inference. Moreover, until comparatively recently there has been the frequent assumption that, if words are decoded, meaning will be automatically understood.

It has long been recognized that reading comprehension is one facet of language communication as a whole. Russell (1965) suggests that the nature of verbal understanding is so complex that researchers frequently become overwhelmed by the enormity of the task involved.

Hunt (in Sheldon, 1964) introduced his study of the nature of comprehension by raising the following questions which are representative of the foundations upon

which studies have been conducted.

1. Does the reader in the usual situation use mental functions or processes which are sufficiently different to reflect a described difference?

2. Do individual readers vary in either their ability or proficiency to use described skills of comprehension to such an extent that we can measure them with existing measures?

3. Can we assemble groups of items which are true measures of the described abilities or skills?

4. Does each item group measure one designated skill in a manner that is significantly different from its power to measure other skills?

Hunt sought to determine whether six factors--word knowledge, ability to manipulate ideas and concepts in relation to one another, ability to grasp the author's expressed ideas, ability to identify the writer's intent or purpose, ability to follow the organization of a passage, and knowledge of literary devices and techniques--made up aspects of reading comprehension, and if they could be measured as independent variables. Hunt developed tests to measure each of the six factors, taking pains to insure that every item included as a part of the test for a given factor was judged with very high consistency, by competent consultants, to be measuring the

ability in question. Using 585 college students, he first compared item-discrimination values for every item with respect to each of the six postulated factors. Despite all preparations, Hunt found that in general the items classified in any given area correlated no higher with the total score on the ability they were supposed to measure than with the score on any of the other abilities. Factor analysis of the scores led to the same general conclusion (Hunt, in Sheldon, 1964).

Traxler (1941) sought to determine whether reading tests evaluated specific aspects of comprehension as a basis for diagnostic work. Studying the results of the VanWagenen-Dvorak Diagnostic Examination of Silent Reading Abilities given to 116 tenth-grade students, he found that the skills of comprehension which were tested were so closely correlated that they nearly became one.

Conant (1942) attempted to discern whether there was a "general reading comprehension" or a number of different reading skills which determined reading proficiency. She administered her own test, containing specific concrete and abstract skills, along with the Nelson-Denny Reading Test and the American Council Psychological Examination. Intercorrelations among all the measures except five were above .50, leading Conant to conclude that there was a single factor which could be labeled "general

comprehension." Allowance was made for the existence of individual differences in the use of different reading techniques.

Only 25 percent or more of the variance was accounted for in Conant's study. This indicates that Conant makes a poor interpretation of weak findings.

At the high school level, Holmes and Singer (1966) defined 8 factors in a matrix of 56 variables, but power of reading, assessed by the VanWagenen-Dvorak Diagnostic Examination of Silent Reading Abilities, correlated significantly with only 3 factors: 66 percent on audio-visual verbal symbolic reasoning, 9 percent on auditory perception, and 3 percent on speed of visual perception.

Harris (1948) identified seven skills or behaviors called for in comprehension of various types of literature, as follows: recognition of synonyms for uncommon words and groups of words; recognition of words or groups of words that are used figuratively; recognition of antecedents of pronouns, subjects, and predicates in loosely organized statements; recognition of summary of ideas expressed or implied; recognition of summaries and characteristics of persons or characters; recognition of author's attitude toward his characters, his mood or emotion, and his intent; and recognition of relationship between technique and meaning.

After administering a battery of tests designed to yield measures on each of these seven skills to two groups of adults and factor-analyzing the results, Harris concluded that: (1) one and only one ability is common to the comprehension of literary passages of different types; and (2) that one general factor is adequate to account for the intercorrelations of the seven variables. Harris obtained very little specificity in his test, but he was mainly concerned with comprehension in literature. His separate tests do not seem to have succeeded in measuring different abilities.

Davis (1944) sought to identify some of the fundamental factors in reading comprehension and to provide a means of measuring them. On the basis of a comprehensive survey of the literature, he listed nine supposed categories of basic skills of reading comprehension. He proceeded to develop test questions to measure each of these skills, administered the tests to a group of subjects, and computed the intercorrelations among the nine tests. He interpreted a factor analysis of the results as indicating the presence of nine factors, six of them clearly significant. These latter included word knowledge, ability to manipulate ideas and concepts in relation to one another--"reasoning in reading," ability to grasp the author's expressed ideas, and ability to identify

the writer's use of literary devices and techniques. Of Davis' nine factors, word knowledge accounted for by far the greatest part of the variance, followed by the so-called "reasoning in reading" and the literal meaning factors.

Davis concluded that at least two factors, the word knowledge and the reasoning factors, were measured in his tests with sufficient reliability for practical use, and that adequately reliable measures of three other factors--literal meaning, inference, and ability to follow the organization of a selection--could be developed as a practical matter.

A reanalysis of Davis' data by Thurstone (1946), employing a somewhat different factor analysis technique, led Thurstone to conclude that a single factor was sufficient to account for the obtained correlations. The apparent conflict in interpretation reflects different purposes served by the respective types of factor analysis employed in the two investigations. Davis (1946), reacting to Thurstone's reanalysis of his data, continued to maintain that his first six factors, at least, represented significant dimensions of reading comprehension, though admittedly several of them accounted for very little variance in reading scores.

Using cross-validating uniqueness analysis, Davis

(1968) further substantiated his earlier conclusions. Making a distinction between the skills used in comprehension by mature readers, Davis concluded that comprehension is not a unitary trait. Memory for word meaning and drawing inferences about the context of passages had the largest unique nonchance variance in the set of eight skills. Three other skills which accounted for appreciable percentages of unique variance were: following the structure of a passage; recognizing a writer's purpose, attitude, tone, and mood; and finding answers to questions asked explicitly or in paraphrase.

A preponderance of the literature in this area suggests that there is a general comprehension ability, but that this factor will not account for all of the total performance. The accountability of the remaining variation lies within other individual differences characterized by varying reading techniques utilized with a variety of materials.

Correlations of Intelligence and Reading Tests

Although psychologists have not been able to agree on a definition of intelligence, the three ideas that occur most frequently in definitions are that it involves ability to deal effectively with abstractions, ability to learn, and ability to respond appropriately in new

situations (English and English, 1958).

The degree of correlation between measured intelligence and reading performance varies with the tests used, as well as with the ages of children.

Individual verbal tests (e.g., Stanford-Binet) tend to correlate with reading success in the neighborhood of .60 to .70. The primary grade IQ tests, which utilize oral directions and have a high verbal content but use pictures in order to avoid reading, tend to have correlation with reading scores about like those of the Stanford-Binet (Traxler and Townsend, 1955).

Verbal group mental ability tests from the fourth grade up tend to have higher correlations with reading scores, ranging from .70 to about .85, while the so-called nonverbal or non-language group tests have much lower correlations with reading tests, generally ranging between .20 and .40 (Traxler and Townsend, 1955).

The intelligence tests which correlate substantially with progress in reading and other school work have been shown to be weighted with cultural factors. On the other hand, tests which come closer to being culture-free have such low correlations with scholastic success as to be of little use for academic prediction (Eels, 1951).

Harris (1948) has suggested that, because most good tests of intelligence and of reading have probable

errors of measurement ranging between 2 and 6 months, an individual's reading ability should be at least 6 months below his intelligence level before one can be reasonably confident that his reading is definitely below expectation.

Results of numerous studies of the relationship of reading achievement and intelligence have led to the conclusion that intelligence is a major factor in reading success at all levels. Analyses by Bond (1938), Bond and Fay (1950), and Strang (1943) show that this relationship becomes increasingly more pronounced as populations are sampled at succeedingly higher grade levels.

Even though intelligence is related to successful achievement in reading, as it is to all other learning, this fact does not necessarily guarantee reading success for the children with a high IQ. Betts (1956) concluded that eight out of ten retarded readers have normal or superior intelligence. Kottmeyer (1959), not as extreme, states that it is not at all uncommon for bright pupils to develop reading disability, although most remedial readers will be dull or normal in intelligence.

The use of intelligence tests for prediction has also been challenged by Harrington and Durrell (1955), since reading difficulties occur among children at virtually all intellectual levels. Consideration must also

be given to the question of whether intelligence tests measure the important perceptual aspects of reading success and failure. In addition, intelligence scores of retarded readers are often spuriously low when measured by a group intelligence test which requires reading.

It is difficult, according to Newman (1969), to secure a dependable measure of intelligence for the extremely poor reader. Often the child's "intelligence" has to pass through many reading skills before it even becomes measurable by means of such an instrument. Experience in the reading clinics has shown that an appreciable number of children referred to as dull--with IQ's of between 80 and 90 on group intelligence tests--have scored as much as 20 to 30 points higher when given individual tests. The individual intelligence test, of course, gives the opportunity to observe responses directly and is, therefore, not only more diagnostic but more reliable. However, when dealing with the poor reader, it is wise to view with reservation and to regard as highly tentative the IQ score derived from any test involving the use of language.

Immediate Recall and Comprehension

The memory or association factor which must occur when ideas are assimilated from language has had scant attention until recently. Yet the ways in which the

recipient receives and processes information he receives may be very pertinent to this study. Miller (1967) has undertaken some basic research here. However, it has yet to be ascertained how far this activity is affected by temperament, by intrinsic styles, and by previous experience and training; or whether there are generalities implicit in both the method and the content of communication which applies to all humans.

The comprehension criterion is especially important in testing oral and silent reading. The criterion is met if items test comprehension and interpretation rather than pure memory of what has been read. Some poor readers can remember very well what they read, but understand little of it.

Barbe (1958) suggests that memory is not a reading skill as such, but a facet of intelligence. A common factor of memory may account for the high correlations which are reported between the group reading tests and group intelligence tests.

The investigator regarded the immediate recall factor as an important influence in the determination of students' performance in reading comprehension tests as used by remedial reading teachers of junior high school students.

A search of the literature revealed a near dearth

of findings in the area of effect of immediate recall (as defined in Chapter I) upon comprehension. No study was found which sought to determine the amount of influence which immediate recall has upon comprehension. This study will attempt to fill that vacuum.

CHAPTER III

PROCEDURES

This chapter will describe the investigation of the results of oral and silent reading tests and intelligence tests, the materials used to instruct and evaluate results, methods of testing, and the treatment of data for significance.

Description of the Sample

The sample for this study was composed of 40 seventh-grade students from Kawameeh Junior High School, Union, New Jersey.

Kawameeh Junior High School is located in a suburban, predominantly white area of Union. Information obtained from school cumulative records showed that the majority of subjects came from families where the father was employed in skilled labor, white- or blue-collar jobs, or self-employed.

The students had been classified as remedial reading students by joint judgment of their seventh-grade teachers and counselors. The only criterion for placement in the class was that each child was reading below his

potential, judgments being based on objective and subjective measures.

Materials Used for the Study

The Trigg Diagnostic Reading Tests, Survey Section: Upper Level; Form A (1966), subtests Story Comprehension and Comprehension, were utilized. The Story Comprehension subtest is composed of story-type material followed by 20 questions used to measure the extent of the student's immediate recall and comprehension of what he has read. Eight minutes are allowed for the reading of the selection, and 7 minutes are given for the answering of questions. The median of reliabilities (Forms A-H) for this score is .72; the average validity for Form A is .47 (Buros, 1968).

The Comprehension subtest is composed of four selections of reading materials similar to those found in textbooks in social studies and science. Each selection is followed by five questions based on the reading materials in the selections. Fifteen minutes are allowed for the reading of the selections and the answering of questions. The reliability given for this subtest is a combination of the two subtests used in this study; a separate reliability for this subtest is not stated. The average validity for Form A is .48 (Buros, 1968).

The Spache Diagnostic Reading Scales (1963), Comprehension section, is composed of 22 reading passages

of graduated difficulty of reading materials that might be found in grades one through eight. They are narrative, expository, and descriptive selections drawn from a variety of sources. The median reliability for this score is .86. Concurrent validity with the California Reading Test is .78; with the Wechsler Intelligence Scale for Children, .80 (Buros, 1968).

The Lorge-Thorndike Intelligence Test (1964) provides both a verbal and nonverbal battery. The verbal battery is made up of five subtests which use only verbal items: vocabulary, verbal classification, sentence completion, arithmetic reasoning, and verbal analogy. The nonverbal battery uses items which are either pictorial or numerical. It contains three subtests involving pictorial classification, pictorial analogy, and numerical relationships. Reliability is reported to be .86. Concurrent validity with the Stanford grade equivalents in reading is .87 (Buros, 1959).

Procedure of Testing

The Spache Diagnostic Reading Scales was administered first to the subjects in one sitting. The test was given to each student individually. The student began testing by orally reading a selection rated a grade below that which the tester thought to be his instructional level. This was done to prevent the child from failing

to succeed on his first selection.

The tester would ask the questions associated with the selection, and the child would respond orally and without a time limit. Success with the reading selection was determined by percentage criteria, which included comprehension and oral reading errors. The subject continued to read up through the graded selections until he could no longer meet the criteria. His last successful level indicated his instructional reading level.

The next selection was read silently, but the questioning procedure remained the same. The last level successfully passed indicated the student's independent reading level.

There was a minimum lapse of 4 days between the administrations of the Spache Diagnostic Reading Scales and the Triggs Diagnostic Reading Test to offset any of many numerous types of influence from the first test upon the second.

The Triggs Diagnostic Reading Test was administered to the subjects in groups of three or four. Total testing time was divided into two sittings--one for each subtest.

The Lorge-Thorndike group IQ test had been administered to the students at the beginning of the seventh grade, 4 months before the study began.

Treatment of the Data

The reading tests and IQ tests described above were given to 40 seventh-grade remedial reading students in Union, New Jersey, during the 1970-1971 school year.

The Triggs Diagnostic Reading Test was hand scored by means of a punched overlay key and the tables provided in the manual. The Spache Reading Scale was also hand scored as per manual direction. The Lorge-Thorndike Intelligence Test scores had been machine graded.

The raw scores for the subjects were keypunched on Fortran cards by the examiner. The data were processed at the Center for Computer and Information Services, Rutgers--The State University of New Jersey. The program used was BMD03D Correlations with Item Deletion. The system card, problem card, F-type variable format card, plot selection card, and finish card were prepared and arranged with the data input cards as outlined in BMD: Biomedical Computer Programs (Dixon, 1967).

Means, standard deviations, and simple correlations are presented in Chapter IV.

The correlations computed by means of the BMD03D program were used to calculate partial correlations of the five variables. This was done to maximize the efficiency of test prediction and to define further the contributing factors of each test.

The formula for calculating the partial correlation coefficient to eliminate the effects of a third variable is:

$$r_{12 \cdot 3} = \frac{r_{12} - r_{13}\, r_{23}}{\sqrt{1 - r_{13}^2}\ \sqrt{1 - r_{23}^2}} \quad \text{(McNemar, 1955)}$$

The investigator limited the factors to be under test of the partial correlation computations to two: verbal comprehension and immediate recall.

Examination of the items and the directions for taking the five tests accounted for the following determination of factor composition:

Triggs Comprehension:	Verbal comprehension
Triggs Story Comprehension:	Verbal comprehension + recall
Spache Silent Reading:	Verbal comprehension + recall
Spache Oral Reading:	Verbal comprehension + recall
Lorge-Thorndike IQ:	Verbal and nonverbal comprehension

Note: The Lorge-Thorndike test was not partialed out of the intercorrelation, even though no immediate recall was necessary, because of the presence of the nonverbal comprehension factor.

The calculation of the partial correlation coefficients yielded the following: the partial correlation coefficient, the variance interpretation of the

proportional overlap among variables, the proportional overlap with a variable eliminated, and the proportional overlap resulting from the effect of a particular variable. The variable chosen by the investigator to be eliminated was Verbal Comprehension, as represented by the Triggs Comprehension subtest. These results are presented in Chapter IV.

CHAPTER IV

RESULTS AND DISCUSSION

This study investigated if different comprehension subtests, as found in the Triggs Diagnostic Reading Test and the Spache Diagnostic Reading Test, measure the same facets of comprehension. More specifically, it sought to find the proportion of the variance accounted for by comprehension test scores if the need for immediate recall of the material is partialed out of the scores. This investigation also attempted to determine the correlation between group intelligence tests with comprehension tests which necessitate immediate recall and those which do not.

Results

Means and standard deviations of the five tests used in the study for the total group (N = 40) are given in Table 1. The Spache Silent Reading Test score (representing the Independent Reading Level) mean of 6.3 was derived from grade-level raw scores of 4.5 to 8.5. The standard deviation was 1.2.

The Spache Oral Reading Test score (representing the Instructional Reading Level) mean of 6.0 was derived from grade-level raw scores of 3.5 to 8.5. The standard

TABLE 1

MEANS AND STANDARD DEVIATIONS OF SPACHE DIAGNOSTIC SCALE, TRIGGS DIAGNOSTIC TEST, AND LORGE-THORNDIKE IQ (IN RAW SCORE FORM)

(N = 40)

Measures	Means	S.D.
Triggs Comprehension	7.3	3.3
Triggs Story Comprehension	6.8	2.3
Spache Silent Reading	6.3	1.2
Spache Oral Reading	6.0	1.4
Lorge-Thorndike IQ	96.5	7.4

deviation was 1.4.

The Triggs Story Comprehension Test score mean of 6.8 was derived from a raw score range of 1 to 11 out of a possible 20 points. The standard deviation was 2.3.

The Triggs Comprehension Test score mean of 7.3 was derived from a raw score range of 0 to 11 out of a possible 20 points. The standard deviation was 3.3.

The Lorge-Thorndike Intelligence Test score mean was 96.5. The standard deviation was 7.4.

Relationships among the five variables tested appear in the correlation matrix in Table 2. Of the ten intercorrelations among the four comprehension subtests and IQ test scores, four are significant at the .05 level or greater. The correlation had to be equal to .267 for this level of significance. Garrett (1966) described a correlation of from ±.40 to ±.70 as usually denoting "substantial or marked relationship" between two variables, and ±.20 to ±.40 as significant but "low."

The major question of this study--do different comprehension tests measure the same facets of comprehension?--can only be answered tentatively.

Examination of Table 2 reveals that the three correlations significant at the .01 level were .50 for Triggs Comprehension and IQ scores, .47 for Spache Oral and Spache Silent Reading Tests, and .39 for Spache Silent

TABLE 2

CORRELATIONS AMONG SPACHE DIAGNOSTIC SCALES, TRIGGS DIAGNOSTIC TEST, AND LORGE-THORNDIKE IQ (IN RAW SCORE FORM)

(N = 40)

	1 Triggs Compre-hension	2 Triggs Story	3 Spache Silent	4 Spache Oral	5 Lorge-Thorndike IQ
2	.36*				
3	.12	.28*			
4	.22	.20	.47**		
5	.50**	.01	.39**	.20	

*p < .05.

**p < .01.

and IQ scores.

Correlations significant at the .05 level were .36 for Triggs Story Comprehension and Triggs Comprehension scores, and .28 for Triggs Story Comprehension and Spache Silent Comprehension scores.

The lowest correlations occurred between: Triggs Story Comprehension and Spache Silent Comprehension scores at .22, Triggs Story Comprehension and Spache Oral Comprehension at .20, Spache Oral Comprehension and IQ scores at .20, Triggs Comprehension and Spache Silent Comprehension at .12, and Triggs Story Comprehension and IQ scores at .01. None of these three correlations were significant at the .05 level.

Table 3 squares the correlations to yield the overlapping variance. It is apparent that only from .0001 to .25, or .01 percent to 25 percent, of the total variance is accounted for in any of the intercorrelations.

The variable in Column 1 (Triggs Comprehension: verbal comprehension) was eliminated by means of partial correlation.

E.g.: The elimination of verbal comprehension from Triggs Story and Spache Silent was computed utilizing the formula:

$$r_{23 \cdot 1} = \frac{r_{23} - r_{12}\, r_{13}}{\sqrt{1 - r_{12}^2}\ \sqrt{1 - r_{13}^2}}$$

TABLE 3

PROPORTIONAL OVERLAP AMONG VARIABLES
IN TERMS OF VARIANCE

	1 Triggs Compre-hension	2 Triggs Story	3 Spache Silent	4 Spache Oral	5 Lorge-Thorndike IQ
2	.13				
3	.01	.08			
4	.05	.04	.22		
5	.25	.0001	.15	.04	

The elimination of the verbal comprehension from Triggs Story and Spache Oral was computed utilizing the formula:

$$r_{24 \cdot 1} = \frac{r_{24} - r_{12}\, r_{14}}{\sqrt{1 - r_{12}^2}\ \sqrt{1 - r_{14}^2}}$$

The same type of computation was carried out for the four partial correlations.

Table 4 shows that the range of the partial correlations was from -.21 to .45. The lowest partial correlation was between IQ and Spache Oral subtests.

Each of the partial correlations was interpreted in variance terms, and each variance proportion was treated for two areas of accountability: verbal comprehension and other factors.

The second question of this study--what proportion of the variance is accounted for in comprehension test scores if the need for immediate recall of the material is partialed of the score--is answered in Tables 5 and 6 which show the proportion of the variance accounted for comprehension test scores if the need for verbal comprehension is partialed out of the score. The variation accounted for by verbal comprehension ranged from -.04 to .03. The amount of variation accounted for by factors other than verbal comprehension ranged from .01 to .20.

TABLE 4

PARTIAL CORRELATION COEFFICIENTS ELIMINATING THE TRIGGS COMPREHENSION TEST (VERBAL COMPREHENSION)

	1 Triggs Compre-hension	2 Triggs Story	3 Spache Silent	4 Spache Oral	5 Lorge-Thorndike IQ
2	X				
3	X	.26			
4	X	.13	.45		
5	X	-.21	.38	.10	

TABLE 5

AMOUNT OF VARIANCE DUE TO THE EFFECT OF VERBAL COMPREHENSION

	1 Triggs Compre-hension	2 Triggs Story	3 Spache Silent	4 Spache Oral	5 Lorge-Thorndike IQ
2	X				
3	X	.02			
4	X	.02	.02		
5	X	-.04	.01	.03	

TABLE 6

AMOUNT OF VARIANCE AFTER THE ELIMINATION OF THE EFFECT OF VERBAL COMPREHENSION

	1 Triggs Compre-hension	2 Triggs Story	3 Spache Silent	4 Spache Oral	5 Lorge-Thorndike IQ
2	X				
3	X	.06			
4	X	.02	.20		
5	X	.04	.14	.01	

This study failed to answer the third question posed by the investigator. The correlation of .50 between Triggs Comprehension scores and IQ scores was significant at the .01 level. This would indicate that comprehension tests and IQ tests not demanding the use of immediate recall correlate fairly well. But this indication is made invalid by the .39 correlation, significant at the .01 level, between IQ scores and the Spache Silent scores which demands immediate recall. Thus, this study failed to consistently answer the question--how highly do group intelligent tests correlate with comprehension scores which necessitate immediate recall and those which do not.

Discussion

The Triggs Comprehension subtest scores and the IQ scores correlated at .50, which was significant at the .01 level. This significant correlation may have been due to the possibility that neither test demanded immediate recall on the part of the subject, who was free to refer back to the material in an attempt to answer the questions.

However, the correlation of .50 between these two tests accounts for only 25 percent of the variance of one variable's predictability from the other, and the residual 75 percent is due to other factors.

Conversely, the significant correlation of .47 between the two subtests of the Spache Diagnostic Test may

have been due in part to the need for immediate recall on the part of the subject in order to answer the question successfully. A partial correlation of these two subtests yielded .45; thus, holding the verbal comprehension constant did not drastically affect the relationship.

The necessity, or lack of necessity, of immediate memory can account for the extremely insignificant correlation (.01) between the IQ test (which does not require immediate recall) and the Triggs Story Comprehension subtest (which does require immediate recall).

This same difference in memory demands may be responsible for the low correlation (.20) between IQ and Spache Oral Comprehension scores. Partialing out verbal comprehension reduced the original low correlation to .10. This shows that verbal comprehension may have a depressing effect on the observed correlation between IQ and Spache Oral.

This reasoning may also be applied in explaining the correlations of .12 between Triggs Comprehension and Spache Silent subtests and .22 between Triggs Comprehension and Spache Oral subtests.

If memory were a unifying factor in the test and subtests under question, one would expect Spache Oral and Triggs Story Comprehension subtests and Spache Silent and Triggs Story Comprehension subtests to be highly

correlated. They were not. This leads the investigator to assume that other factors of comprehension contributed to a lack of unity. Verbal comprehension did contribute 50 percent to the cumulative variance of Spache Oral and Triggs Story subtests and 25 percent of the cumulative variance of Spache Silent and Triggs Story subtests. These percentages indicate the substantial influence of the excluded variable to these subtests.

Triggs Story (immediate recall required) and Triggs Comprehension subtests (no immediate recall) significantly correlated; but to a degree so that only 12 percent of the variance was accounted for. The investigator assumed that immediate recall would account for a large proportion of the remaining 88 percent.

The correlation between the IQ test scores and the Triggs Comprehension scores has the highest of the ten raw correlations. The investigator assumes that this is because both tests allow the student to peruse the selection freely and often in the discrimination process of choosing an answer.

In contrast to this are two of the other correlations involving IQ, those with Triggs Story Comprehension and Spache Oral Comprehension subtests. The two correlations were at less than the .05 significance level. Difference in memory demands may be partially responsible for

this insignificant correlation, but other variables, not under test, account for the greatest part of the variance.

Both of these comparisons involved a test requiring heavy recall and a test requiring no recall. Partialing out the verbal comprehension influence, the correlations arrived at were negative. Thus, it appears that the students possessing high verbal comprehension may have less, or need less, immediate recall.

Although four of the correlations were significant at the .05 level or better, very little of the variances of any of the five variables could be explained by, attributed to, or be predictive of the variance of any other variable.

Interpretation of the partial correlations revealed that a very small amount of the total variances of the five variables was determined by the verbal comprehension factor. Indeed, there was a negative relationship between verbal comprehension and unexplained influences in some of the partial correlations.

Only 1 percent to 22 percent of the variation in comprehension scores (raw correlations) was accounted for by the Triggs and Spache subtests.

The range of accountability involving IQ scores was a bit wider, from .0001 percent to 25 percent.

The remaining 78 to 99+ percent of the variances

may lie in factors which were not tapped or did not emerge in this study. But the outstanding common denominator of immediate recall in three of the five variables led the investigator to conclude that immediate recall does account for a large (but as yet undetermined) part of the unexplained variation as far as these particular tests are concerned.

It must be remembered when considering the results of this investigation that the tests chosen for study were selected not only for their testing of comprehension but also in their demands of immediate recall. Therefore, the investigator can only disagree with the findings of Conant (1942) or Traxler (1941), insofar as they did not qualify the particular circumstances under which their conclusions were valid. Their studies of comprehension did not involve immediate recall situations; thus their conclusion of the existence of a "general comprehension" cannot be fully counted out.

The results of the study tend to add some support to Barbe's (1958) assertion that memory may account for the high correlations which occur between some reading and intelligence tests.

The substantial to low correlations between IQ and reading comprehension tests are consistent with Newman's (1969) and Harrington and Durrell's (1955) contentions

that it is wise to view the poor readers' IQ scores with reservation in light of the amount of language involved in the test's usage.

However, none of the correlations of the present study were above .50; none of the partial correlations were above .25. This does not negate the possibility that there may be a limited number of factors which account for comprehension. But this limited number of factors which account for the greater portion of variability under the conditions of this study are more than one in number.

CHAPTER V

SUMMARY AND CONCLUSIONS

This investigation studied the composition of reading comprehension. More specifically, it was an attempt to understand some relationships among the comprehension subtests of oral reading tests, silent reading tests, and intelligence tests by means of inter- and partial correlations.

This investigation studied the role of immediate recall in oral and silent reading comprehension tests and intelligence tests. More specifically, the investigation concerned the relationships between comprehension subtests of the Spache Diagnostic Scales, the Triggs Diagnostic Reading Tests, and the Lorge-Thorndike Intelligence Test.

Forty seventh grade remedial reading students participated in the study. The three tests were administered to the students in their remedial reading class periods.

Raw scores were processed at the Center for Computer and Information Services, Rutgers--The State University, on the IBM 7040 using BMD03D Correlations with Item Deletion. Means, standard deviations, and intercorrelations of the variables were obtained for the total sample. Partial correlations were hand calculated.

Conclusions

The results of this study indicate that the validities usually attributed to the tests utilized in this investigation must be reconsidered when measuring the reading level and status of junior high school remedial reading students.

The findings have supported the investigator's hypothesis that readers may use differing comprehension factors in seemingly similar circumstances. One of these variables--such as verbal comprehension--can be measured independently, and does account for some of the predictability of the response. The study has shown that tests requiring a great deal of immediate recall are interrelated, though not to a high degree of significance. Likewise, those subtests which demand less use of immediate memory and more rationalization are more highly intercorrelated. Memory, though, does not account for all of the difference among the measures.

Holding the verbal comprehension factor constant produced no drastic change in any of the original intercorrelations. This indicates that verbal comprehension has little effect on any of the observed correlations. The nature of the remaining subtests suggests that a large part of remaining relationships among these subtests may be due to immediate recall. Perhaps, since most of the

variance was not accounted for by partial correlations, variables other than those under consideration are responsible for the greater part of the observed associations.

For this study it can be concluded that comprehension is not, on the whole, general; and, although there may be some overlap in abilities, different facets or skills are applied in various seemingly similar tests.

The components of the reading process are represented in different proportions in the tests in varying degree of effectiveness. The reading teacher must examine the actual test items to get a clear understanding of what skills and abilities the test is measuring; it is only as he so examines that he can judge whether the given test is valid for his purposes.

The findings of this investigation imply that two reputable tests with comprehension subtests are evaluating very different skills of comprehension with little overlap.

Suggestions for Further Study

Since this study did not produce any highly significant relationships among any of the variables, or account for any large proportion of the variance, incorporation of a number of changes in the research design are recommended in any replication of the study. These recommendations would include:

1. A balance between reading tests requiring

memory and those which do not (this study had an imbalance of one subtest of verbal comprehension and three of verbal comprehension and immediate recall);

2. Tests of comprehension which measure specific factors of comprehension, which would allow for more specificity in variance accountability;

3. Computation of the results in standard score form, rather than raw scores, so that comparisons of the means and standard deviations between tests can be made; and

4. Use of intelligence tests which measure verbal comprehension only.

BIBLIOGRAPHY

Anderson I., & Swanson, D. Common factors in eye-movements in silent and oral reading. Psychological Monographs, 1937, 43(3), 61-79.

Barbe, W. B. Measuring reading comprehension. The Clearing House, 1958, 32, 343-345.

Betts, E. A. Are retarded readers dumb? Education, 1956, 37, 568-575.

Betts, E. A. Foundations of reading instruction. New York: American Book Co., 1957.

Bond, E. Reading and ninth grade achievement. New York: Bureau of Publications, Teachers College, Columbia University, 1938.

Bond, E., & Fay, L. C. A comparison of the performance of good and poor readers on the individual items of the S-B Scales, Form L and M. Journal of Educational Research, 1950, 43, 457-459.

Botel, M. Newer techniques in the supervision of reading instruction. In Frontiers of elementary education. Syracuse, N. Y.: Syracuse University Press, 1957.

Buros, O. K. (Ed.) The fifth mental measurements yearbook. Highland Park, N. J.: Gryphon Press, 1959.

Buros, O. K. (Ed.) Tests and reviews. Highland Park, N. J.: Gryphon Press, 1968.

Buswell, G. An experimental study of the eye-voice span in reading. Supplementary Educational Monographs, No. 17. Chicago: University of Chicago Press, 1920.

Chall, J. Interpretation of results of standardized reading tests. Supplementary Monographs, No. 88. Chicago: University of Chicago Press, 1958.

Cole, L. The improvement of reading. New York: Holt, Rinehart and Winston, 1938.

Committee on Diagnostic Reading Tests, Inc. Diagnostic reading tests. Mountain Home, N. C.: n.p., 1961.

Conant, M. The construction of a diagnostic reading test. New York: Columbia University, Teachers College, 1942.

Davis, F. B. Fundamental factors of comprehension in reading. Psychometrika, 1944, 9, 185-197.

Davis, F. B. A brief comment on Thurstone's note on a reanalysis of Davis' reading tests. Psychometrika, 1946, 11, 249-255.

Davis, F. B. Identification and measurement of reading skills of high school students. Cooperative research project No. 3023. Philadelphia: University of Pennsylvania, 1967.

Davis, F. B. Research in comprehension in reading. Reading Research Quarterly, 1968, 3, 499-545.

Dixon, W. I. (Ed.) BMD: Biomedical computer programs. Berkeley: University of California Press, 1967.

Dolch, E. W. Diagnostic and remediation of disability in oral reading. Supplementary Educational Monographs, No. 82. Chicago: University of Chicago Press, 1955.

Eells, K. Intelligence and cultural differences. Chicago: University of Chicago Press, 1951.

English, H. B., & English, A. C. A comprehensive dictionary of psychological and psychoanalytical terms. New York: Longmans, Green, 1958.

Fairbanks, G. The relationship between eye movements and voice in the oral reading of good and poor silent readers. Psychological Monographs, No. 3. Chicago: University of Chicago Press, 1937.

Garrett, H. E. Statistics in psychology and education. New York: David McKay, 1966.

Gates, A. The improvement of reading. New York: Macmillan, 1947.

Harrington, Sr. M. J., & Durrell, D. D. Mental maturity vs. perception abilities in primary reading. Journal of Educational Psychology, 1955, 46, 375-380.

Harris, A. J. Diagnosis of reading disabilities. Supplementary Educational Monographs, No. 79. Chicago: University of Chicago Press, 1953. (a)

Harris, A. J. Motivating the poor reader. Education, 1953, 73, 566-574. (b)

Harris, A. J. How to increase reading ability. (3rd and 4th eds.) New York: Longmans, Green, 1956.

Harris, C. W. Measurement of comprehension of literature: II. Studies of measures of comprehension. School Review, 1948, 56, 332-342.

Hildreth, G. Reading program in the early primary period. In N. Henry (Ed.), 48th Yearbook of National Society for the Study of Education--Part II: Reading in the elementary school. Chicago: University of Chicago Press, 1949.

Holmes, J. A., & Singer, H. Speed and power of reading in high school. U.S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Monograph No. 14. Washington, D.C.: U.S. Government Printing Office, Superintendent of Documents Catalog No. FS 5.230:30016, 1966.

Julitta, Sr. M. Selection and use of standardized reading tests. Supplementary Educational Monographs, No. 88. Chicago: University of Chicago Press, 1957.

Killgallon, P. A. A study of relationships among certain pupil adjustments in reading situations. As cited in E. A. Betts, Foundations of reading instruction. New York: American Book Co., 1957.

Kottmeyer, W. Teacher's guide for remedial reading. New York: Webster Publishing, 1959.

McCracken, R. A. The development and validation of the standard reading inventory for the individual appraisal of reading through classroom practice. International Association Conference Proceedings, No. 9. Newark, Del.: International Reading Association, 1964.

McCullough, F. What's behind the reading score? Elementary English, 1953, 30, 1-7.

McNemar, Q. Psychological statistics. New York: John Wiley and Sons, 1955.

Miller, G. A. The psychology of communication. New York: Basic Books, 1967.

Newman, H. (Ed.) Reading disabilities. New York: Odyssey Press, 1969.

O'Brien, I. A. Silent reading. New York: Macmillan, 1921.

Russell, D. H. Research on the process of thinking with some applications to reading. Elementary English, 1965, 42, 370-378.

Schmidt, W. A. An experimental study in the psychology of reading. Chicago: University of Chicago Press, 1917.

Sheldon, W. What is reading comprehension? A research view. In A. Mazurkiewicz (Ed.), New perspectives in reading instruction. New York: Pitman Publishers, 1964.

Spache, G. Diagnostic reading scales. Monterey: California Test Bureau, 1963.

Strang, R. Relationships between certain aspects of Intelligence and certain aspects of reading. Educational and Psychological Measurement, 1943, 3, 355-359.

Swanson, D. E. Common elements in silent and oral reading. Psychological Monographs, 1937, 48(3), 36-60.

Thorndike, E. L. Reading as reasoning: A study of mistakes in paragraph reading. Journal of Educational Psychology, 1917, 8, 323-332.

Thurstone, L. L. Note on a reanalysis of Davis' reading tests. Psychometrika, 1946, 11, 185-188.

Traxler, A. A study of the Van Wagenen-Dvorak diagnostic tests: Examination of silent reading abilities. Educational Records Bulletin, 1941, No. 31, pp. 33-41.

Traxler, A., & Townsend, A. Eight more years of research in reading. Educational Records Bulletin, 1955, No. 64, pp. 51-63.

APPENDIX C

RAW DATA SCORES

Raw Data Scores

	Triggs Subtest		Spache Diagnostic		Lorge-Thorndike
	Comprehension	Story	Silent	Oral	IQ
1	13	11	6.5	6.5	94
2	10	8	6.5	6.5	111
3	11	8	4.5	4.5	100
4	3	1	6.5	5.5	97
5	7	7	7.5	6.5	101
6	8	6	7.5	7.5	110
7	2	6	6.5	5.5	94
8	8	7	6.5	6.5	113
9	7	9	6.5	6.5	91
10	11	6	6.5	6.5	103
11	9	8	7.5	6.5	97
12	10	9	7.5	7.5	96
13	6	5	4.5	4.5	83
14	7	11	7.5	8.5	88
15	8	4	4.5	3.5	93
16	3	3	4.5	3.5	96
17	8	4	8.5	7.0	104
18	7	8	8.5	5.0	97
19	9	4	6.5	6.0	107
20	11	8	5.5	6.0	102
21	9	8	5.5	3.0	92
22	2	5	6.5	6.0	92
23	1	2	6.5	6.0	88
24	7	4	6.5	5.0	95
25	7	5	4.5	3.0	84
26	6	10	7.5	7.0	95
27	7	6	4.5	7.5	89
28	11	10	6.5	7.5	104
29	11	9	8.5	5.5	108
30	102	6	6.5	7.5	102
31	109	5	5.5	3.5	109
32	94	7	5.5	3.5	94
33	87	8	5.5	5.5	87
34	95	8	6.5	5.5	95
35	90	8	6.5	5.5	90
36	101	7	6.5	6.5	101
37	94	8	4.5	4.5	94
38	97	7	6.5	6.5	97
39	99	9	5.5	5.5	99
40	92	6	4.5	4.5	92